An open and closed case for all polymerases

Sylvie Doublié[1], Michael R Sawaya[2] and Tom Ellenberger[2]*

The recently determined structures of HIV-1 reverse transcriptase and Taq DNA polymerase in complex with DNA primer–template and an incoming nucleotide have shown that a large conformational change configures the polymerase active site for nucleotidyl transfer. The structure of reverse transcriptase in the catalytic complex will open the path to the rational design of novel nucleoside analog inhibitors of viral replication.

Addresses: [1]Department of Microbiology and Molecular Genetics, The Markey Center for Molecular Genetics, University of Vermont, Burlington, VT 05405, USA and [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115-5730, USA.

*Corresponding author.
E-mail: tome@jeckle.med.harvard.edu

Structure February 1999, 7:R31–R35
http://biomednet.com/elecref/09692126007R0031

Introduction

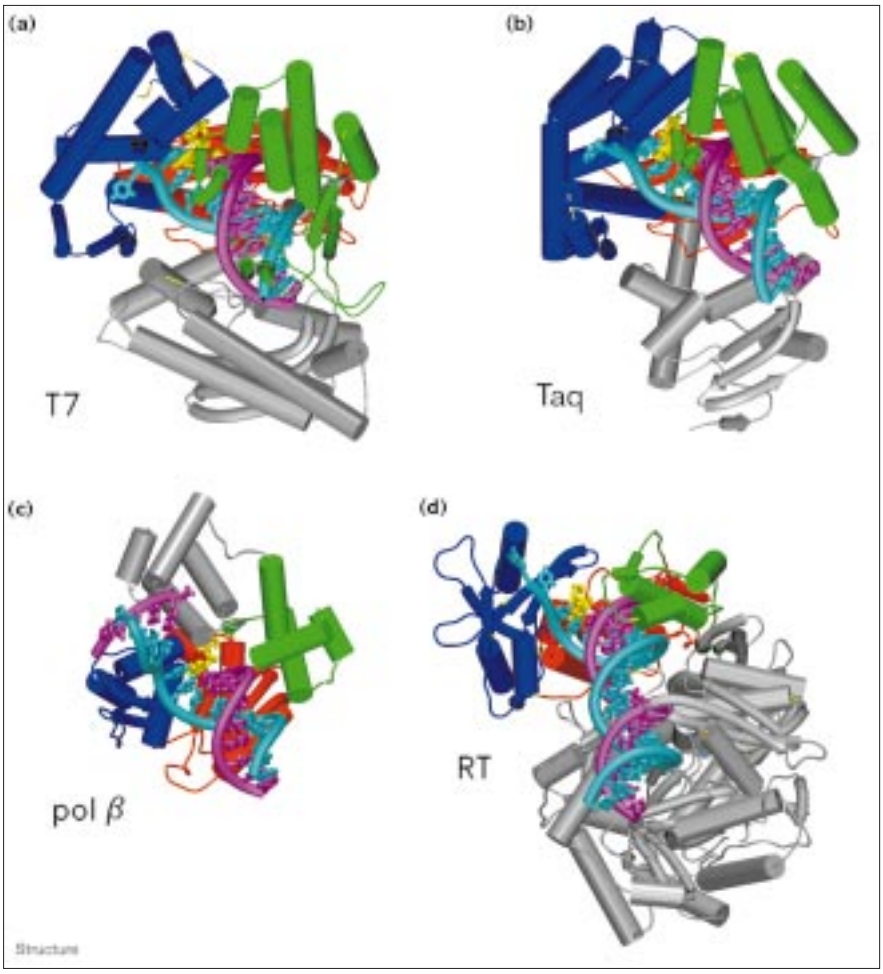
DNA polymerases faithfully copy a DNA template by the stepwise addition of nucleotides onto the terminal 3′-OH of a DNA primer. These enzymes have vastly different sizes, cellular roles, replication fidelities, and sensitivities to nucleotide analogs. Until recently, it appeared that the sole unifying theme among all DNA polymerases was a common chemistry for nucleotidyl transf er that utilizes bound divalent metal ions and several clusters of conserved amino acid residues [1]. Polymerases sharing more extensive sequence homology have been grouped into four families (pol I, pol α, polymerase family X, and reverse transcriptase). The past few years have borne a bounty of exciting structural and biochemical findings for each of the four families of DNA polymerase [2–15]. The structural portraits of polymerases from all four families have revealed similar silhouettes, featuring a U-shaped DNA-binding cleft that resembles a partially opened right hand with 'fingers', 'thumb', and 'palm' subdomains, as was originally described for the large fragment of *Escherichia coli* DNA polymerase I (Klenow fragment) [16]. The detailed secondary structures represented by the four DNA polymerase families are dissimilar, however. Despite their structural diversity, the emerging data suggest that all DNA polymerases employ analogous catalytic schemes for template-directed DNA synthesis. This common strategy for DNA synthesis is evidenced by the similar geometry of the bound DNA, nucleotide, and metal ions seen in ternary complexes of mammalian polymerase β (pol β) [9–11], T7 DNA polymerase [7], and more recently, the *Thermus aquaticus* (Taq) DNA polymerase [17] and human immunodeficiency virus type 1 reverse transcriptase (HIV-1 RT) [18]. All of these structures feature a conserved active-site geometry that includes a kink in the 5′-end of the template strand and, most strikingly, a large conformational change in the fingers in response to substrate binding.

The making of a ternary complex

DNA polymerases copy a DNA template with high fidelity, as any mistake would jeopardize the stable transmission of genetic information. In order to study the structural basis for accurate DNA synthesis, it was necessary to capture a polymerase in the act of adding a nucleotide to the 3′-end of the primer strand. The first snapshot of a ternary complex of a polymerase, a DNA primer–template, and an incoming nucleotide was obtained for the mammalian repair enzyme pol β [9]. Following the incorporation of a chain terminating 2′,3′-dideoxynucleotide, pol β was crystallized in complex with a bound nucleotide paired to a template base in the polymerase active site. Although the 3′-end of the DNA primer is poised for nucleophilic attack of the nucleotide α-phosphate, the primer lacks a 3′-OH group and the polymerase is therefore unable to incorporate the bound nucleotide into DNA. This strategy, pioneered for pol β, was employed in subsequent crystallographic studies of T7 DNA polymerase [7], Taq DNA polymerase [17], and HIV-1 RT [18].

A number of X-ray structures of HIV-1 RT have been determined, including those of the unliganded enzyme [12], several complexes with small-molecule inhibitors [14], and a binary complex with a DNA primer–template [13]. The crystallization of HIV-1 RT complexed to both DNA and nucleotide substrates has proven to be an elusive goal — until now. Huang *et al.* devised a clever combinatorial disulfide cross-linking strategy that stabilized the HIV-1 RT–DNA complex in a specific register after several cycles of polymerization [18]. A model based on the previously determined structure of HIV-1 RT bound to a DNA primer–template [13] indicated that amino acid sidechains on the exposed face of helix H of the thumb would be in close contact with the minor groove of the DNA primer–template during synthesis [19]. Three mutant polymerases were prepared with cysteine residues individually substituted at each of the exposed helix H positions. These mutant enzymes were paired with DNA templates containing a thiol-bearing linker at the N2 position of specific guanines in the sequence. Incubation of each one of the HIV-1 RT cysteine mutants with the thiol-modified DNA led to a discrete number of polymerization cycles, followed by the incorporation of a 2′,3′-dideoxynucleotide and

Figure 1



The structures of DNA polymerase ternary complexes. (a) T7 DNA polymerase, (b) Taq DNA polymerase, (c) pol β, and (d) HIV-1 RT in complex with DNA (shown in magenta and cyan) and nucleotide (yellow) substrates. The fingers subdomain is shown in dark blue, the thumb in green, and the palm in red. All four DNA polymerases adopt a closed conformation upon binding DNA and the correct incoming nucleotide. For clarity, only the p66 subunit of HIV-1 RT is shown, without its RNase H C-terminal domain, and the processivity factor thioredoxin was omitted in the T7 DNA polymerase model.

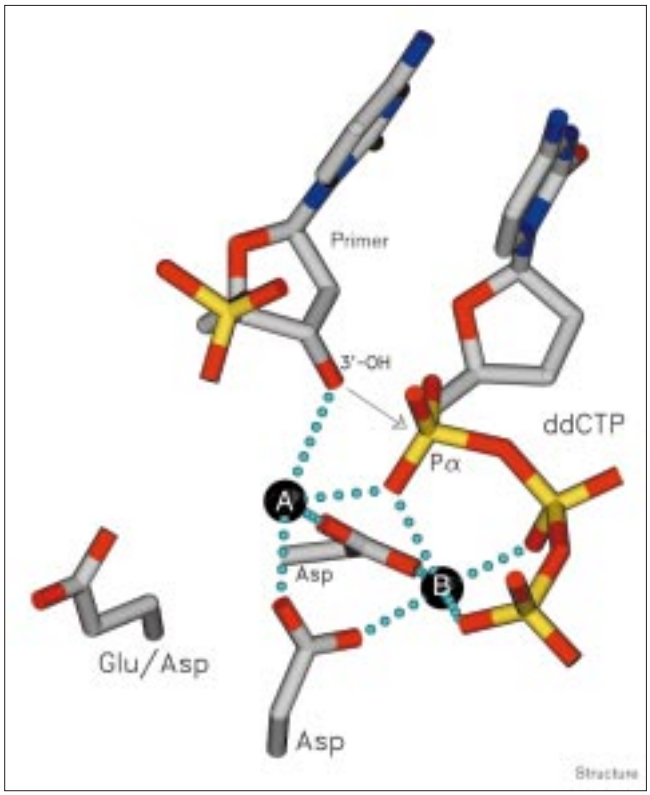
covalent trapping of the disulfide-linked polymerase–DNA complex. A site-specific cross-link only forms when the polymerase is stalled by a chain-terminating nucleoside analog, strongly implying that the resulting complex closely resembles the catalytic state. The disulfide cross-link seems to exert little, if any, effect on the structure of the complex. This cross-linking strategy may therefore be adapted for crystallographic studies of other protein–DNA complexes.

A common mechanism for all DNA polymerases

The structures of four DNA polymerases from three different families have now been determined in complex with DNA and nucleotide substrates (Figure 1) [7,11,17,18]. There is very little similarity between the amino acid sequences of these polymerases, except for a few acidic residues located in the palm subdomain [1]. Despite this very limited sequence identity and differences in overall structure, the three families of enzyme seem to have adopted a common mechanism for nucleotidyl transfer (Figure 2). There are two metal ions (usually magnesium ions) in the active site that are ligated by the conserved acidic residues of the palm. In all four ternary complexes, one metal ion contacts all three phosphates of the bound nucleotide and the other metal ion is situated between the α-phosphate of the nucleotide and the 3′-end of the primer. The conservation of metal-binding sites in these highly divergent DNA polymerases underscores the importance of the metal ions for assisting nucleotide polymerization. One metal ion activates the 3′-OH of the DNA primer for attack of the nucleotide α-phosphate, while the other metal ion stabilizes the displaced pyrophosphate moiety. This mechanism is similar to that first proposed for the 3′→5′ proofreading exonuclease of the Klenow fragment of *E. coli* DNA polymerase I [20].

In all four polymerase complexes, the B-form DNA structure becomes A-form near the polymerase active site. The widened minor groove of the primer–template is contacted by active-site residues that assist in positioning the 3′-end of the primer for the nucleotidyl transfer reaction. The DNA template strand does not go through the cleft formed by the fingers and thumb subdomains. Instead,

Figure 2



Schematic of a common nucleotidyl transfer mechanism for DNA polymerases. Two magnesium ions in the polymerase active site (A and B) are contacted by two strictly conserved aspartate residues. Both metal ions ligate the α-phosphate of the incoming nucleotide, whereas metal ion B also ligates the β- and γ-phosphates. The 3′-OH of the primer terminus is positioned for an in-line attack on the α-phosphate, followed by the release of pyrophosphate. Atoms are shown in standard colors; the metal ion coordination is represented as dotted lines.

there is a sharp kink at the 5′-end of the template that directs the single-stranded DNA template out of the polymerase active site. The template wraps around the fingers subdomain, allowing the DNA polymerase to check the shape of the first base pair in the active site.

A structural basis for the high fidelity of DNA polymerization

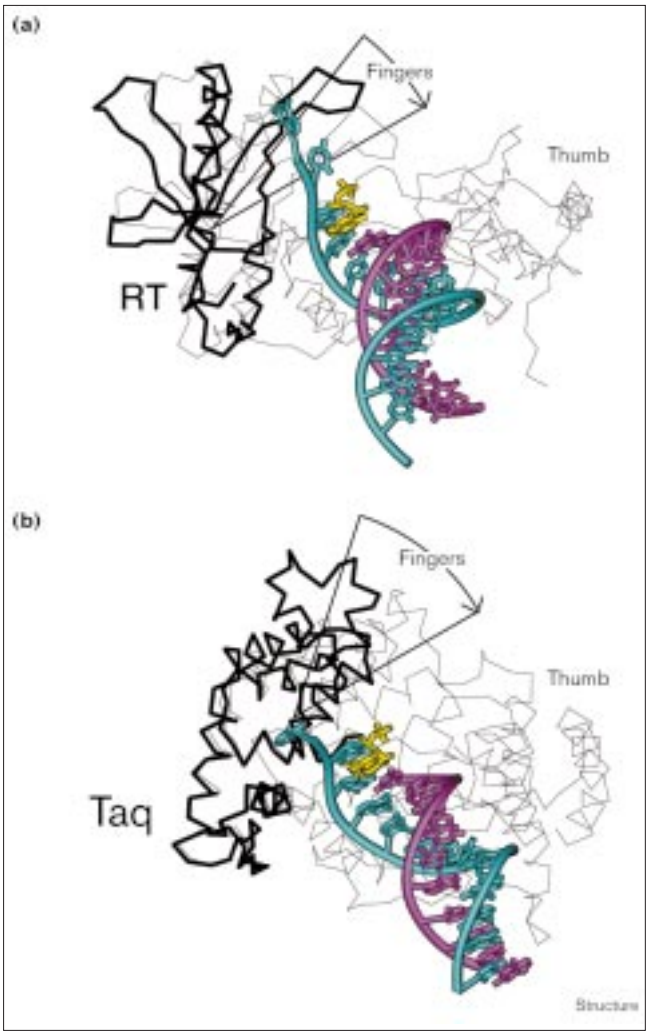
A comparison of the structure of pol β complexed to DNA and a nucleotide with structures of unliganded pol β or of the enzyme bound only to DNA shows that a C-terminal domain closes around the correct incoming nucleotide, thus assembling the polymerase active site for nucleotide incorporation [9–11]. In this closed conformation, pol β can check the geometry of the nucleotide–template base pair in the polymerase active site. This variation of an induced-fit mechanism specifies the correct nucleotide substrate. The ternary complex of T7 DNA polymerase also showed a closed conformation of the polymerase active site, caused by the inward rotation of the fingers subdomain towards the palm [7] (SD and TEE, unpublished results). In a crystallographic tour de force, a series of structures of the Taq polymerase, with and without bound DNA and nucleotide substrates, have directly demonstrated that the fingers subdomain moves by a combined hinge-like motion and a repacking of several helices within the fingers, closing the active site in the catalytic complex (Figure 3) [17]. Interestingly, the Taq binary complex described in this paper was obtained by simply diffusing the nucleotide out of the crystals during crystal harvest — the crystal-packing arrangement allowed the fingers to open, resulting in the formation of a binary complex. The open conformation of the fingers in the Taq binary complex [4,17] is very similar to that observed for the unliganded Taq DNA polymerase [2,3]. The corresponding domain in HIV-1 RT, the β3–β4 loop, is not helical but shows an analogous closure upon binding the correct nucleotide: the fingertips bend inwards towards the palm by about 20° (Figure 3) [18].

The emerging conclusion from crystal structures of a variety of otherwise unrelated DNA polymerases is that the assembly of a catalytic complex with nucleotide and DNA is accompanied by a large conformational change in the protein. The closed conformation of the fingers subdomain allows conserved, functionally important residues to contact the sugar and phosphate moieties of the incoming nucleotide. Kinetic analyses of enzymatic DNA synthesis have identified a rate-limiting step following the binding of a nucleotide substrate and before its incorporation into DNA [21,22]. This slow step might correspond to the closing of the fingers around the bound substrates. The resulting tight fit of the nascent base pair within the polymerase active site precludes a mismatch between template and nucleotide. A distortion caused by a mismatched base pair would prevent assembly of the active-site conformation and presumably lead to the reopening of the fingers and the sampling of another candidate nucleotide.

Drug resistance and the importance of a ternary complex for HIV-1 RT

Reverse transcriptase is a lynch pin for the replication of the HIV-1 virus and, thus, is an important target for antiviral therapies. Most antiviral agents that are currently in clinical use against HIV-1 infection are chain terminating nucleoside analogs that inhibit HIV-1 RT, such as 3′-azido-2′,3′-dideoxythymidine (AZT), 2′,3′-dideoxyinosine (ddI), and 2′-deoxy, 3′-thiacytidine (3TC). The structure of the HIV-1 RT catalytic complex determined by Huang *et al.* is the first ternary complex of HIV-1 RT. In contrast to previous structures of HIV-1 RT lacking nucleotide and/or DNA, the structure of the polymerase–DNA–dNTP complex shows that most of the residues associated with resistance to nucleoside inhibitors contribute directly to the nucleotide-binding site of the polymerase [18]. Other residues having secondary or reinforcing effects on the level of drug resistance are located in the periphery of the

Figure 3



Closure of the fingers upon formation of a polymerase catalytic complex. Open (shown in thick lines) and closed (thin lines) conformations in (a) HIV-1 RT and (b) Taq DNA polymerase. The tip of the fingers subdomain of Taq DNA polymerase rotates by 46° inwards towards the active site when the enzyme is complexed with DNA (shown in magenta and cyan) and nucleotide (yellow) substrate. The corresponding domain in HIV-1 RT, the β3–β4 loop, bends inwards towards the palm by about 20°. Only the fingers subdomain is shown for the open conformations of both polymerases.

nucleotide-binding site. This long awaited view of the HIV-1 RT active site will undoubtedly aid in the design of new nucleoside analogs that are sterically compatible with the enzyme active site, yet unlikely to be rejected by the residue substitutions associated with resistance to the present therapeutic agents.

Concluding remarks

The recently determined crystal structures of DNA polymerases bound to DNA and nucleotide substrates have provided many insights into the mechanism of action of these enzymes. Despite the many differences in the structures of the four DNA polymerase ternary complexes, several universal features are evident. In all cases, the fingers have closed around substrates bound to the polymerase active site. This change in protein conformation brings conserved residues in contact with the incoming nucleotide and aligns it for a backside attack by the activated 3′-OH of the DNA primer. Two metal ions in the active site assist in this reaction by orienting the reacting molecules and dissipating the negative charge that develops during nucleotidyl transfer. In the closed conformation, the polymerase active site snugly accommodates an undistorted template–nucleotide base pair. It is the exquisite fit of the incipient base pair in the active site that ensures that the correct nucleotide is incorporated. The fingers must then reopen so that the DNA template can translocate to the next position in the sequence and a new nucleotide substrate can bind.

Acknowledgements

We thank Stephen Harrison, Gregor Verdine, and Gabriel Waksman for providing information and structural coordinates prior to publication. Our work on DNA replication proteins is supported by the National Institutes of Health, the Armenise-Harvard Center for Advanced Scientific Research, and the Harvard Center for Structural Biology.

References

1. Delarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. (1990). An attempt to unify the structure of polymerases. *Protein Eng.* **3**, 461-467.
2. Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. & Waksman, G. (1995). Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5-Å resolution: structural basis for thermostability. *Proc. Natl Acad. Sci. USA* **92**, 9264-9268.
3. Kim, Y., Eom, S.H., Wang, J., Lee, D.S., Suh, S.W. & Steitz, T.A. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* **376**, 612-616.
4. Eom, S.H., Wang, J. & Steitz, T.A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. *Nature* **382**, 278-281.
5. Kiefer, J.R., *et al.*, & Beese, L.S. (1997). Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution. *Structure* **5**, 95-108.
6. Kiefer, J.R., Mao, C., Braman, J.C. & Beese, L.S. (1998). Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**, 304-307.
7. Doublié, S., Tabor, S., Long, A.M., Richardson, C.C. & Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251-258.
8. Wang, J., Sattar, A.K., Wang, C.C., Karam, J.D., Konigsberg, W.H. & Steitz, T.A. (1997). Crystal structure of a Pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087-1099.
9. Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. & Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**, 1891-1903.
10. Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. & Kraut, J. (1994). Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science* **264**, 1930-1935.
11. Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J. & Pelletier, H. (1997). Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205-11215.
12. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783-1790.
13. Jacobo-Molina, A., *et al.*, & Arnold, E. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl Acad. Sci. USA* **90**, 6320-6324.

14. Ren, J.S., *et al*., & Stammers, D. (1995). High resolution structures of HIV-1 RT from four RT–inhibitor complexes. *Nat. Struct. Biol.* 2, 293-302.
15. Georgiadis, M.M., Jessen, S.M., Ogata, C.M., Telesnitsky, A., Goff, S.P. & Hendrickson, W.A. (1995). Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukemia virus reverse transcriptase. *Structure* 3, 879-892.
16. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. & Steitz, T.A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 313, 762-766.
17. Li, Y., Korolev, S. & Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.* 17, 7514-7525.
18. Huang, H., Chopra, R., Verdine, G.L. & Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 282, 1669-1675.
19. Bebenek, K., *et al.*, & Kunkel, T.A. (1997). A minor groove binding track in reverse transcriptase. *Nat. Struct. Biol.* 4, 194-197.
20. Beese, L.S. & Steitz, T.A. (1991). Structural basis for the 3′→5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* 10, 25-33.
21. Wong, I., Patel, S.S. & Johnson, K.A. (1991). An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry* 30, 526-537.
22. Spence, R.A., Kati, W.M., Anderson, K.S. & Johnson, K.A. (1995). Mechanism of inhibition of HIV-1 reverse transcriptase by non nucleoside inhibitors. *Science* 267, 988-993.